AD-A088 247 MARYLAND UNIV COLLEGE PARK DEPT OF MATHEMATICS F/6 12/1

ON NEAREST NEIGHBOR DEGENERACIES OF INDISTINGUISHABLE PARTICLES--ETC(U)

JUN 80 B KEDEM F49620-79-C-0095

UNCLASSIFIED MD80-37-BK AFOSR-TR-80-0589 NL

1 of 1

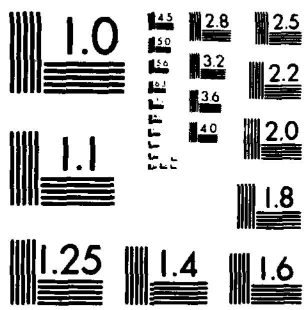
AD A088247

END

DATE FILMED

9-80

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LEVEL II

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-80-0589
2. GOVT ACCESSION NO.: AD-A088247
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): On Nearest Neighbor Degeneracies of Indistinguishable Particles.
5. TYPE OF REPORT & PERIOD COVERED: Interim rept.
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Benjamin Kedem
   MD80-37-EK, TR80-35
8. CONTRACT OR GRANT NUMBER(s): F49620-79-C-0095
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Maryland, College Park, MD 20742
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research /NM, Bolling AFB, Washington, DC 20332
12. REPORT DATE: June 1980
13. NUMBER OF PAGES: 18
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 22
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC ELECTE AUG 26 1980 S D E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Counting, Binary Sequences, Runs, Units, Sufficient Statistic, Poisson, Interaction Energy, Markov

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Arrangement degeneracies suggested by sufficient statistics associated with binary stationary m'th order Markov chains are discussed, and are shown to correspond and generalize some degeneracies arising when indistinguishable particles are placed on a one-dimensional lattice with n compartments. From these statistics it is possible to define an m-th order unit. The arrangement degeneracy obtained from s 1's and n-s 0's so that lower order units are placed in higher order units is solved. However, the most general arrangement degeneracy associated with all the sufficient statistics of a given order is difficult. For this

AD A088247

DDC FILE COPY

DD FORM 1473 400953 UNCLASSIFIED

20. (Continued)

case only the 3rd order arrangement degeneracy is obtained, the 1st and 2nd orders being relatively simple. These results are applied in determining the asymptotic distributions of rare events.

Accession For	
NTIS GRA&I	✓
DDC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Codes	
Dist.	Avail and/or special
A	

AFOSR-TR- 80-0589

ON NEAREST NEIGHBOR DEGENERACIES OF INDISTINGUISHABLE PARTICLES

by

Benjamin Kedem *

Department of Mathematics
University of Maryland
College Park, Maryland 20742

MD80-37-BK
TR80-35
June 1980

*Research supported by AFOSR F49620-79-C-0095 Approved for public release; distribution unlimited.

80 8 14 120

ON NEAREST NEIGHBOR DEGENERACIES OF

INDISTINGUISHABLE PARTICLES

Benjamin Kedem

ABSTRACT

Arrangement degeneracies suggested by sufficient statistics associated with binary stationary m'th order Markov chains are discussed, and are shown to correspond and generalize some degeneracies arising when indistinguishable particles are placed on a one-dimensional lattice with n compartments. From these statistics it is possible to define an m-th order unit. The arrangement degeneracy obtained from s 1's and $n-s$ 0's so that lower order units are placed in higher order units is solved. However, the most general arrangement degeneracy associated with all the sufficient statistics of a given order is difficult. For this case only the 3rd order arrangement degeneracy is obtained, the 1st and 2nd orders being relatively simple. These results are applied in determining the asymptotic distributions of rare events.

ON NEAREST NEIGHBOR DEGENERACIES OF INDISTINGUISHABLE PARTICLES

Benjamin Kedem

INTRODUCTION: In statistical mechanical treatment of cooperative phenomena such as magnetic spin, binary alloys, elasticity, etc., the total energy of interaction E_i is given as a linear combination of potential energies associated with certain arrangements of indistinguishable particles on a one dimensional lattice. For example if only occupied nearest neighbors and next nearest neighbors of the types 11, 101, 111, are of interest then

$$E_i = n_{11}V_{11} + n_{101}V_{101} + n_{111}V_{111} \tag{1}$$

where the V's stand for potential energy and the n's refer to the frequency of occurence of the neighbor type. Here 1 refers to an occupied site while 0 refers to a vacancy. Of interest then is knowledge of the arrangement degeneracy associated with the types of neighbors. That is, the number of binary sequences of size n which satisfy certain restrictions (numbers of types). Much attention to these combinatorial problems has been given by McQuistan in a series of articles in which he considered both simple and complex particles such as dumbbells, e.g. see references [3][4][5]. Parallel to the physical interpretation of arrangement degeneracies there is a purely statistical one in which arrangement degeneracies are used in approximating important distributions such as the distributions of crossings and upcrossings of a fixed level by a stationary process and the distribution of extremes in such processes. This has recently been dealt with in [2].

The purpose of this article is to show the connection between the purely statistical and statistical mechanical approaches in regard to arrangement degeneracies by examining sufficient statistics associated with m'th order Markov chains. At the same time we shall extend some of McQuistan results by introducing higher order degeneracies whose usefulness will be demonstrated in finding the asymptotic distribution of "rare" events.

2. A CONNECTION BETWEEN NEAREST AND NEXT NEAREST NEIGHBOR DEGENERACY AND SUFFICIENCY

Let $\{X_t, t=0,\pm1,\ldots\}$ be a two state (0-1) stationary Markov chain and consider a binary time series from the chain, $X_1,\ldots,X_n$. The sufficient statistics associated with the chain are [2]

$$S = \sum_{i=1}^{n} X_i, \quad R_1 = \sum_{i=2}^{n} X_i X_{i-1}, \quad H = X_1 + X_n.$$

In order to find the joint distribution of S, R_1, H it is necessary to determine the number of binary sequences $M_n(s,r_1,h)$ for which $S = s$, $R_1 = r_1$, $H = h$. This number is easily found to be

$$M_n(s,r_1,h) = \binom{2}{h}\binom{s-1}{r_1}\binom{n-s-1}{s-r_1-h}, \tag{2}$$

and the number of sequences for which only the first two conditions are satisfied is

$$M_n(s,r_1) = \sum_{h=0}^{2} M_n(s,r_1,h) = \binom{s-1}{r_1}\binom{n-s+1}{s-r_1}, \qquad (3)$$

a result which was also obtained by McQuistan [4] for the degeneracy of nearest neighbor pairs.

If the chain $\{X_t\}$ is of 2^{nd} order then [2] the sufficient statistics are S, R_1, H together with $R_2 = \sum_{i=3}^{n} X_i X_{i-2}$, $C = \sum_{i=3}^{n} X_i X_{i-1} X_{i-2}$, $U = X_2 + X_{n-1}$, $V = X_1X_2 + X_{n-1}X_n$. Again the joint distribution of these statistics requires the knowledge of the number of binary sequences $M_n(s,r_1,r_2,c,h,u,v)$ for which $S = s,\ldots,V = v$. This was found [1] to be

$$\begin{aligned} M_n(s,r_1,r_2,c,h,u,v) &= \binom{2}{\max(h,u)}\binom{\max(h,u)}{v}\binom{r_1-1}{c}\binom{s-r_1-h}{r_1-c-v} \\ &\quad \cdot \binom{s-r_1-1}{r_2-c}\binom{n-2s+r_1+h-2}{s-r_1-r_2+c-h-u+v}, \end{aligned} \qquad (4)$$

with the convention $\binom{-1}{-1} = 1$ and where (h,u,v) takes values in $\{(0,0,0),(0,1,0),(0,2,0),(1,0,0),(2,0,0),(1,1,0),(1,1,1),(2,1,1),(1,2,1),(2,2,2)\}$. It follows that the number of binary sequences for which only the first four conditions are fixed is obtained by summing over (h,u,v). We have

$$M_n(s,r_1,r_2,c) = \sum_{(h,u,v)} M_n(s,r_1,r_2,c,h,u,v)$$

$$= \sum_{(0,0,0),(0,1,0),(0,2,0)} + \sum_{(1,0,0),(2,0,0),(1,1,0)}$$

$$+ \sum_{(1,1,1),(2,1,1),(1,2,1)} + M_n(s,r_1,r_2,c,2,2,2)$$

$$= \binom{r_1-1}{c}\binom{s-r_1-1}{r_2-c}\left[\binom{s-r_1}{r_1-c}\binom{n-2s+r_1}{s-r_1-r_2+c} + 2\binom{s-r_1-1}{r_1-c}\binom{n-2s+r_1}{s-r_1-r_2+c-1}\right.$$

$$+ \binom{n-2s+r_1}{s-r_1-r_2+c-2}\binom{s-r_1-2}{r_1-c} + 2\binom{s-r_1-1}{r_1-c-1}\binom{n-2s+r_1}{s-r_1-r_2+c-1}$$

$$\left. + 2\binom{n-2s+r_1}{s-r_1-r_2+c-2}\binom{s-r_1-2}{r_1-c-1} + \binom{n-2s+r_1}{s-r_1-r_2+c-2}\binom{s-r_1-2}{r_1-c-2}\right]$$

$$= \binom{r_1-1}{c}\binom{s-r_1-1}{r_2-c}\binom{s-r_1}{r_1-c}\left[\binom{n-2s+r_1}{s-r_1-r_2+c} + 2\binom{n-2s+r_1}{s-r_1-r_2+c-1} + \binom{n-2s-r_1}{s-r_1-r_2+c-2}\right]$$

$$= \binom{r_1-1}{c}\binom{s-r_1}{r_1-c}\binom{s-r_1-1}{r_2-c}\binom{n-2s+r_1+2}{s-r_1-r_2+c}. \tag{5}$$

Upon noting that $r_1 = n_{11}$, $c = n_{111}$, $r_2-c = n_{101}$, where $n_{ij\ldots k}$ refers to the frequency of $ij,\ldots,k$ in the binary sequence, we recognize the next nearest neighbor degeneracy obtained by McQuistan [4]

$$M_n(s,n_{11},n_{101},n_{111}) = \binom{n_{11}-1}{n_{111}}\binom{s-n_{11}}{n_{11}-n_{111}}\binom{s-n_{11}-1}{n_{101}}\binom{n-2s+n_{11}+2}{s-n_{11}-n_{101}}. \quad (5')$$

Obviously (5') can be obtained from (5) by linearity. Here, and in what follows, we use the notation $M_n(c_1,\ldots,c_k)$ to denote the number of binary sequences for which $c_1,\ldots,c_k$ are fixed. It is readily seen that (5) reduces to (3) by summing over r_2-c and c.

3. HIGHER ORDER DEGENERACIES

In the previous section we illustrated via two examples that arrangement degeneracies may be viewed as special cases of counting problems associated with sufficient statistics in Markov chains. The next thing which comes to mind is the question of generalizations. It is quite clear now that if one wants to look into or define higher degeneracies one should examine the sufficient statistics and linear functions thereof of higher order two state Markov chains. By the very definition of sufficient statistics, it is intuitively clear that every conceivable arrangement degeneracy can be obtained by summing over M_n of a given order. We need not consider all the sufficient statistics associated with a given order but only those which define desired neighbor types. To make this point clear we shall defer the general counting problem to the next section while concentrating here on straightforward generalizations of (3) and (5). We shall first illustrate this claim by examining sufficient statistics associated with 3^{rd} order chains. In this connection the notion of a "unit" is useful.

DEFINITION: An m'th order unit is a binary sequence which starts with a 1, ends with m separating 0's, (if needed to separate it from other units) and in which each 0-run, if not an end run, consists of at most $m-1$ 0's. Thus a unit is a block made of 0-runs and 1-runs where the length of the 0-runs is restricted while the length of the 1-runs is unrestricted.

DEFINITION: A free 0 is a 0 which does not belong to any unit.

Observe that an end unit is a unit which does not need the separating 0's at the end in order to be recognized. The notion of a unit is useful as it determines the general form of a binary series which satisfies our predetermined conditions. For example, consider a first order chain and suppose it is desired to count the number of sequences for which S and R_1 are fixed at s and r_1. There are $s-r_1$ first order units which we permute with the free 0's in $\binom{s-r_1+(n-s)-(s-r_1-1)}{s-r_1} = \binom{n-s+1}{s-r_1}$ ways. Next we place a 0 at the end of each first order unit in one way. This determines the form of the binary series. Finally distribute the s 1's in the $s-r_1$ units so that none is empty in $\binom{s-1}{r_1}$ ways. This yields (3). (5) can be obtained in the same way by first determining the positions of the 2^{nd} order units and then placing the 1^{st} order units in the 2^{nd} order units. This is followed by the distribution of the s 1's. This procedure gives rise to an immediate extension of (3) and (5).

Consider the 3^{rd} order analog of (3) and (5). The statistics of interest are S, R_1, R_2, C, together with $K = \sum_{i=4}^{n} X_i X_{i-1} X_{i-2} X_{i-3}$ and n_{1001}, the frequency of 1001 in the sequence. These are sufficient statistics (not all of them!) associated with 3^{rd} order chains. We shall determine $\underline{M}_n(s,r_1,r_2,c,k,n_{1001})$.

There are $(s-r_1) - (r_2-c) - n_{1001}$ 3^{rd} order units, $(s-r_1) - (r_2-c)$ 2^{nd} order units and $s-r_1$ 1^{st} order units. First permute the 3^{rd} order units with the free 0's in $\binom{n-3s+2r_1+r_2-c+3}{s-r_1-r_2+c-n_{1001}}$ ways. Next place the 2^{nd} order units in the 3^{rd} order units leaving none empty in $\binom{s-r_1-r_2+c-1}{n_{1001}}$ ways. Then place the 1^{st} order units in the 2^{nd} order units in $\binom{s-r_1-1}{r_2-c}$ ways and place the separating 0's in the respective units in one way. Put one 1 in each 1^{st} order unit in one way. There remain r_1 1's. There are r_1-c 1^{st} order units with two or more 1's. So choose r_1-c 1^{st} order units from $s-r_1$ and put one 1 in each in $\binom{s-r_1}{r_1-c}$ ways. There are $c-k$ 1^{st} order units with 3 or more 1's. So select $c-k$ 1^{st} order units from r_1-c and put 1 in each in $\binom{r_1-c}{c-k}$ ways. Finally place the remaining k 1's in the $c-k$ 1^{st} order units allowing "empty" units in $\binom{c-1}{k}$ ways. Whence

$$M_n(s,r_1,r_2,c,k,n_{1001}) = \binom{c-1}{k}\binom{r_1-c}{c-k}\binom{s-r_1}{r_1-c}\binom{s-r_1-1}{r_2-c}$$

$$\cdot \binom{s-r_1-r_2+c-1}{n_{1001}}\binom{n-3s+2r_1+r_2-c+3}{s-r_1-r_2+c-n_{1001}}, \tag{6}$$

and by summing over n_{1001} we immediately obtain

$$M_n(s,r_1,r_2,c,k) = \binom{c-1}{k}\binom{r_1-c}{c-k}\binom{s-r_1}{r_1-c}\binom{s-r_1-1}{r_2-c}\binom{n-2s+r_1+2}{s-r_1-r_2+c}.$$

This last expression yields (5) upon summation over k. We can rewrite (6) in a form which resembles (5'). Let fz_3 denote the number of free 0's associated with a third order degeneracy. That is, the number of 0's which do not belong to any 3rd order unit. Then $fz_3 = (n-s) - 3[(n-s)-n_{101}-n_{1001}-1] - 2n_{1001} - n_{101}$, and (6) becomes

$$M_n(s,n_{11},n_{111},n_{1111},n_{101},n_{1001}) = \binom{n_{111}-1}{n_{1111}}\binom{n_{11}-n_{111}}{n_{111}-n_{1111}}\binom{s-n_{11}}{n_{11}-n_{111}}$$

$$\cdot \binom{s-n_{11}-1}{n_{101}}\binom{s-n_{11}-n_{101}-1}{n_{1001}}\binom{fz_3+s-n_{11}-n_{101}-n_{1001}}{s-n_{11}-n_{101}-n_{1001}} \tag{6'}$$

From (3), (5'), (6') we see that a pattern of arrangement degeneracies begins to emerge whereby the highest order units are permuted with the free 0's then lower order units are placed in higher order units, successively, then the separating 0's are placed in one way, this followed by the distributions of 1's so that $s, n_{11}, n_{111}, \dots, n_{101}, \dots, n_{100\dots1}$, are preserved. In order to arrive at the general result suggested by the above scheme we shall employ the notation $*(m)$ to mean a $*$-run of length m. Then we readily have

$$M_n(s, n_{11}, n_{111}, \dots, n_{1(m+1)}, n_{101}, n_{1001}, \dots, n_{10(m-1)1}) =$$

$$\binom{n_{1(m)}-1}{n_{1(m+1)}}\binom{n_{1(m-1)}-n_{1(m)}}{n_{1(m)}-n_{1(m+1)}}\cdots\binom{n_{11}-n_{111}}{n_{111}-n_{1111}}\binom{s-n_{11}}{n_{11}-n_{111}}$$

$$\cdot\binom{s-n_{11}-1}{n_{101}}\binom{s-n_{11}-n_{101}-1}{n_{1001}}\cdots\binom{s-n_{11}-n_{101}-\cdots-n_{10(m-2)1}-1}{n_{10(m-1)1}}$$

$$\cdot\binom{fz_m+s-n_{11}-n_{101}-\cdots-n_{10(m-1)1}}{s-n_{11}-n_{101}-\cdots-n_{10(m-1)1}}, \tag{7}$$

where fz_m stands for the number of free 0's and is given by

$$fz_m = (n-s) - m[(s-n_{11})-n_{101}-n_{1001}-\cdots-n_{10(m-1)1}-1]$$
$$- n_{101}-2n_{1001}-\cdots-(m-1)n_{10(m-1)1}. \tag{8}$$

Observe that the number of m'th order units is equal to

$$(s-n_{11}) - n_{101} - n_{1001} - \cdots - n_{10(m-1)1} \tag{9}$$

since we essentially evaluate $\sum X_i(1-X_{i-1})\cdots(1-X_{i-m})$. From (7) we obtain (3), (5'), (6') as special cases.

4. THE GENERAL ARRANGEMENT DEGENERACY OF MARKOV CHAINS AND ITS APPLICATION TO THE DISTRIBUTION OF RARE EVENTS.

In the previous section we dealt with one way of extending the next nearest neighbor degeneracy. However, this is only a special case of the general arrangement degeneracy associated with an m'th order Markov chain where the problem is to count the number of binary sequences for which all the sufficient statistics are fixed. This is a challenging problem for which no general solution exists as far as the present author knows. The main difficulty is the fact that not all the low order units can simply be placed in higher order units as in the previous section since some conditions are violated. Clearly, if such a result is available, numerous arrangement degeneracies can be deduced from it. This solution will not be attempted here. However, we shall give the solution for 3^{rd} order chains and this will give us a clue as to the general behavior of rare events in binary Markov chains.

The sufficient statistics associated with a stationary 3^{rd} order Markov chain, apart from ends statistics, are S, R_1, R_2, C, K as above together with

$$R_3 = \sum_{i=4}^{n} X_i X_{i-3}, \quad R_{12} = \sum_{i=4}^{n} X_i X_{i-1} X_{i-3}, \quad R_{21} = \sum_{i=4}^{n} X_i X_{i-2} X_{i-3}.$$

We shall construct a sequence for which S, $R_1, \ldots, R_{21}$ are fixed. First observe that

$$R_3 - R_{12} - R_{21} + K = n_{1001}. \qquad (10)$$

Also define the statistics

$A_{11-11} \equiv$ # of 2^{nd} order units which start and end with 11.

$A_{1-1} \equiv$ # of 2^{nd} order units which start and end with 1.

$A_{11-1} \equiv$ # of 2^{nd} order units which start with 11 and end with 1.

$A_{1-11} \equiv$ # of 2^{nd} order units which start with 1 and end with 11.

$A_{11} \equiv$ # of 2^{nd} order units which contain exactly two consecutive 1's.

$A_{1} \equiv$ # of 2^{nd} order units which contain exactly one 1.

Clearly

$$A_{11-11} + A_{11-1} + A_{1-11} + A_{1-1} + A_{11} + A_1 = (S-R_1) - (R_2-C)$$

$$A_{11} + A_{11-11} + A_{11-1} = (R_1-C) - (R_{12}-K) \qquad (11)$$

$$A_{11} + A_{11-11} + A_{1-11} = (R_1-C) - (R_{21}-K),$$

which means that if A_1, A_{11}, and A_{11-11} are known then so are A_{1-11}, A_{11-1} and A_{1-1}. As it turns out our problem is simplified greatly if A_1, A_{11}, A_{11-11} are added to the other eight conditions. The reason for introducing the A's is that it is difficult to keep track of R_{12} and R_{21}, while from (11) it is seen that when the A's are fixed in addition to S, R_1, R_2, C, K, so are R_{12}, R_{21}. When the arrangement degeneracy is obtained the A's can be removed by summation.

Recall that there are

$$(s-r_1) - (r_2-c) - (r_3-r_{12}-r_{21}+k)$$

3^{rd} order units and

$$(n-s) - 3[(s-r_1)-(r_2-c)-(r_3-r_{12}-r_{21}+k)-1] - 2(r_3-r_{12}-r_{21}+k) - (r_2-c)$$

free 0's which we permute in

$$\binom{(n-s)-2(s-r_1)+(r_2-c)+3}{(s-r_1)-(r_2-c)-(r_3-r_{12}-r_{21}+k)}$$

ways. There are $(s-r_1) - (r_2-c)$ 2^{nd} order units which we place in the 3^{rd} order units in

$$\binom{(s-r_1)-(r_2-c)-1}{r_3-r_{12}-r_{21}+k}$$

ways. Place the separating 0's in one way. Now, we <u>cannot</u> place the 1^{st} order units in the 2^{nd} order units as we did in the previous section as R_{12}, R_{21} change. This is precisely why we need the A's. So according to A_{11}, A_1, A_{11-1}, A_{1-11}, A_{11-11}, A_{1-1}, assign "types" to the 2^{nd} order units. That is let A_{11-11} 2^{nd} order units start and end with 11, A_{11-1} 2^{nd} order units start with 11 and end with 1, etc. This can be done in

$$\binom{(s-r_1)-(r_2-c)}{A_1,A_{11},A_{11-11},A_{11-1},A_{1-11},A_{1-1}}$$

ways. There are now r_2-c 0's left which we distribute in $A_{11-11} + A_{11-1} + A_{1-11} + A_{1-1}$ 2^{nd} order units "nonempty" in

$$\binom{r_2-c-1}{A_{11-11}+A_{11-1}+A_{1-11}+A_{1-1}-1} = \binom{r_2-c-1}{(s-r_1)-(r_2-c)-(A_{11}+A_1)-1}$$

ways. This takes care of the 0's and the 1^{st} order units as well! It remains now to distribute the 1's. According to the type assignment place 11 and 1 in the 2^{nd} order units as needed in one way, and then put 1 in every empty 1^{st} order unit (1-cell or 1-run) in one way. There are r_1-c 1^{st} order units with two or more 1's. But we have already $A_{11-1} + A_{1-11} + 2A_{11-11} + A_{11}$ 1^{st} order units with two 1's. So select $(r_1-c) - (A_{11-1}+A_{1-11}+2A_{11-11}+A_{11})$ 1^{st} order units from $(s-r_1) - (A_{11-1}+A_{1-11}+2A_{11-11}+A_{11}) - (A_{11-1}+A_{1-11}+2A_{1-1}+A_1)$ 1^{st} order units in

$$\binom{(s-r_1) - (A_{11-1}+A_{1-11}+2A_{11-11}+A_{11}) - (A_{11-1}-A_{1-11}+2A_{1-1}+A_1)}{(r_1-c) - (A_{11-1}+A_{1-11}+2A_{11-11}+A_{11})}$$

ways and put one 1 in each. We now have r_1-c 1^{st} order units with exactly two 1's, since previously no 1^{st} order unit was empty. There are $c-k$ 1^{st} order units with 3 or more 1's. From the r_1-c 1^{st} order units which contain exactly 2 1's select $c-k$ units and put one 1 in each in

$$\binom{r_1-c}{c-k} .$$

ways. We now have $c-k$ 1^{st} order units with exactly 3 1's. Finally, there remain k 1's which we place in these $c-k$ 1^{st} order units allowing "empty" units in

$$\binom{c-k+k-1}{c-k-1} = \binom{c-1}{k}$$

ways. It follows that

$$M_n(s,r_1,r_2,c,r_3,r_{12},r_{21},k,A_{11},A_1,A_{11-11}) = \binom{c-1}{k}\binom{r_1-c}{c-k}$$

$$\cdot \binom{(s-r_1)-2(A_{11-11}+A_{11-1}+A_{1-11}+A_{1-1})-(A_{11}+A_1)}{(r_1-c)-(A_{11-1}+A_{1-11}+2A_{11-11}+A_{11}} \cdot \binom{r_2-c-1}{(s-r_1)-(r_2-c)-(A_{11}+A_1)-1}$$

$$\cdot \binom{(s-r_1)-(r_2-c)}{A_1,A_{11},A_{11-11},r_1-c-r_{12}+k-A_{11}-A_{11-11},r_1-c-r_{21}+k-A_{11}-A_{11-11},A_{1-1}}$$

$$\cdot \binom{(s-r_1)-(r_2-c)-1}{r_3-r_{12}-r_{21}+k} \cdot \binom{(n-s)-2(s-r_1)+(r_2-c)+3}{(s-r_1)-(r_2-c)-(r_3-r_{12}-r_{21}+k)}, \qquad (12)$$

where A_{1-1} is obtained from (11) in terms of A_{11-11}, A_{11}, A_1, s, r_1, r_2, c, r_{12}, r_{21}, k, and is too long to write. Thus the desired arrangement degeneracy is obtained by summing over A_{11}, A_{11-11}, A_1, adhering to the convention $\binom{-1}{-1} = 1$.

We have

$$M_n(s,r_1,r_2,c,r_3,r_{12},r_{21},k) = \binom{c-1}{k}\binom{r_1-c}{c-k}\binom{(s-r_1)-(r_2-c)-1}{r_3-r_{12}-r_{21}+k}$$

$$\cdot \binom{(n-s)-2(s-r_1)+(r_2-c)+3}{(s-r_1)-(r_2-c)-(r_3-r_{12}-r_{21}+k} \tag{13}$$

$$\sum_{A_{11},A_1,A_{11-11}} B_n(s,r_1,r_2,c,r_3,r_{12},r_{21},k,A_{11},A_1,A_{11-11})$$

where B_n is equal to the product of all the coeeficients in (12) which involve A_{11}, A_1, A_{11-11}. Note that

$$\sum_{A_{11},A_1,A_{11-11}} B_n(s,0,0,\ldots,0,A_{11},A_1,A_{11-11}) = 1 \tag{14}$$

since then $A_1 = s$ and the rest of the A's vanish.

Now the joint distribution of a binary time series from a 3^{rd} order stationary Markov chain is given by

$$p(x_1,x_2,\ldots,x_n) = (\text{powers in } x_1,x_2,x_3,x_{n-2},x_{n-1},x_n)$$

$$\cdot\, p_{1011}^{r_{21}-k}\, p_{1101}^{r_{12}-k}\, p_{0101}^{r_2-c-r_{12}+k}\, p_{1010}^{r_2-c-r_{21}+k}\, p_{1001}^{r_3-r_{21}-r_{12}+k}\, p_{1111}^{k}\, p_{1110}^{c-k}\, p_{0111}^{c-k}$$

$$\cdot\, p_{1100}^{r_1-c-r_{12}+k}\, p_{0011}^{r_1-c-r_{21}+k}\, p_{0110}^{r_1-2c+k}\, [p_{1000}p_{0001}]^{s-r_1-r_2+c-r_3+r_{12}+r_{21}-k}$$

$$\cdot\, p_{0100}^{s-2r_1-r_2+2c+r_{12}-k}\, p_{0010}^{s-2r_1-r_2+2c+r_{21}-k}$$

$$\cdot\, p_{0000}^{n-4s+3r_1+2(r_2-c)+r_3-r_{12}-r_{21}+k-3}\,, \tag{15}$$

where $p_{x_i x_{i-1} x_{i-2} x_{i-3}} = \Pr(X_i = x_i | X_{i-1} = x_{i-1}, X_{i-2} = x_{i-2}, X_{i-3} = x_{i-3})$, from which it follows that the joint distribution of S, R_1, R_2, C, R_3, R_{12}, R_{21}, K, X_1, X_2, X_3, X_{n-2}, X_{n-1}, X_n is the product ot (15) and the arrangement degeneracy of these statistics. But when X_1, X_2, X_3, X_{n-2}, X_{n-1}, X_n are equal to 0 this arrangement degeneracy is asymptotically, as $n \to \infty$, the same as (13). Therefore from (13), (14) and (15)

$$\Pr(S=s, R_1=0, R_2=0, \cdots, K=0, X_1=X_2=X_3=X_{n-2}=X_{n-1}=X_n=0)$$

$$\sim \binom{n-3s+3}{s} p_{1000}^s \, p_{0001}^s \, p_{0100}^s \, p_{0010}^s \, p_{0000}^{n-4s-3} \tag{16}$$

Assume that as $n \to \infty$

(i) The 1's become rare separated by long 0-runs so that the event $\{R_1=0, R_2=0, \cdots, K=0, X_1=X_2=X_3=X_{n-2}=X_{n-1}=X_n=0\}$ becomes a sure event.

(ii) $p_{1000} \to 0$ such that $np_{1000} = \alpha$, fixed

(iii) $p_{0001} \sim p_{0100} \sim p_{0010} \sim p_{0000} = 1 - p_{1000}$.

Then from (16), as n becomes large,

$$\Pr(S=s) \sim \frac{(n-3s+3)!}{(n-4s+3)!\,s!} \left(\frac{\alpha}{n}\right)^s \left(1-\frac{\alpha}{n}\right)^{n-s-3}$$

$$= \frac{\alpha^s}{s!} \left(1-\frac{3s-3}{n}\right) \cdots \left(1-\frac{4s-4}{n}\right) \left(1-\frac{\alpha}{n}\right)^{-s-3} \left(1-\frac{\alpha}{n}\right)^n$$

$$\longrightarrow \frac{e^{-\alpha}\alpha^s}{s!}\,, \tag{17}$$

a result which is well expected [2].

This procedure for finding the asymptotic distribution of S as $n \to \infty$ can be easily extended to the m'th order case. Arguing as above we have from (7)(as the last binomial coeeficient in the most general case is given by the last coeeficient in (7) which stands for the number of permutations of the m'th order units with the free 0's)

$$Pr(S=s) \sim \binom{n-ms+m}{s} p_{10(m)}^{s} p_{0(m+1)}^{n-s-m}$$

$$\sim \frac{\alpha^s}{s!} \frac{(n-ms)\cdots(n-(m+1)s+1)}{n^s} \left(1-\frac{\alpha}{n}\right)^n \longrightarrow \frac{\alpha^s}{s!} e^{-\alpha}, \tag{18}$$

where $n \to \infty$ and $p_{100\cdots0} \to 0$ such that $np_{100\cdots0} = \alpha$.

More results of this nature can be obtained once the arrangement degeneracy of a specific order is known. For example, from (12) it should not be too difficult to show that under some conditions similar to (i)-(iii) $S-R_1$ is also asymptotically poisson. In fact poisson with parameter $\alpha(1-\lambda)$ where $\lambda \in (0,1)$ is a measure of the density, or clutering tendency, of rare events. For an interpretation of this fact see [2].

REFERENCES

[1] B. Kedem, Discrete Mathematics, 18, 285(1977).

[2] B, Kedem, Binary Time Series (Marcel Dekker, New York, 1980).

[3] R.B. McQuistan, J. Math. Phys. 15, 1192 (1974).

[4] R.B. McQuistan, J. Math. Phys. 15, 1845 (1974).

[5] R.B. McQuistan, J. Math. Phys. 18, 1191 (1977).

ACKNOWLEDGMENT

The author wishes to thank Don Zagier for his useful remarks and interest in this problem.